

**APPEAL BRIEF****IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re Application of:	§	Group Art Unit: 2176
William K. Bodin, <i>et al.</i>	§	
	§	Examiner: Rutledge, Amelia L.
Serial No.: 10/756,159	§	
	§	Atty Docket No.: AUS920030998US1
Filed: 1/13/2004	§	
	§	Customer No.: 34533
Title: Differential Dynamic Content	§	
Delivery with Text Display in	§	Confirmation No.: 3344
Dependence Upon Sound Level	§	

**Mail Stop: Appeal Brief-Patents**

Commissioner for Patents

P.O. Box 1450

Alexandria, Virginia 22313-1450

**APPEAL BRIEF****Honorable Commissioner:**

This is an Appeal Brief filed pursuant to 37 CFR § 41.37 in response to the Final Office Action of August 31, 2006 (hereafter “the Office Action”) and pursuant to the Notice of Appeal filed October 27, 2006.

**REAL PARTY IN INTEREST**

The real party in interest in accordance with 37 CFR § 41.37(c)(1)(i) is the patent assignee, International Business Machines Corporation (“IBM”), a New York corporation having a place of business at Armonk, New York 10504.

**RELATED APPEALS AND INTERFERENCES**

There are no related appeals or interferences within the meaning of 37 CFR § 41.37(c)(1)(ii).

**STATUS OF CLAIMS**

Status of claims in accordance with 37 CFR § 41.37(c)(1)(iii): Thirty-six claims are filed in the original application in this case. Claims 37 and 38 were added in the Amendment filed June 12, 2006. Claims 1-38 are rejected in the Final Office Action. Claims 1-38 are on appeal.

**STATUS OF AMENDMENTS**

Status of amendments in accordance with 37 CFR § 41.37(c)(1)(iv): No amendments were submitted after final rejection. The claims as currently presented are included in the Appendix of Claims that accompanies this Appeal Brief.

**SUMMARY OF CLAIMED SUBJECT MATTER**

Appellants provide the following summary of the claimed subject matter according to 37 CFR § 41.37(c)(1)(v). This summary includes a concise explanation of the subject matter defined in each of the independent claims involved in the appeal and includes references to Appellants' original specification by page and line number and to the drawings by reference characters. The three independent claims involved in this appeal are claims 1, 13, and 25. Claim 1 recites a computer-implemented method for differential dynamic content delivery. Claims 13 and 25 recite counterpart aspects of the method of claim 1. Claim 13 recites systems aspects of the method of claim 1. Claim 25 recites computer program product aspects of the method of claim 1.

Claim 1 recites a computer-implemented method for differential dynamic content

delivery (page 49, lines 27-29, and Figure 19). The method of claim 1 includes providing a session document for a presentation (page 49, lines 27-29, and Figure 19, elements 450 and 266), wherein the session document includes a session grammar and a session structured document (page 49, line 29 - page 50, line 3, and Figure 19, elements 266, 258, 256, and 450). The method of claim 1 also includes selecting from the session structured document a classified structural element in dependence upon user classifications of a user participant in the presentation (page 50, lines 5-7, and Figure 19, elements 452, and 465). The method of claim 1 also includes presenting the selected structural element to the user (page 50, lines 14-15, and Figure 19, elements 454 and 456). The method of claim 1 also includes streaming speech to the user from one or more users participating in the presentation (page 50, lines 24-27, and Figure 19, elements 568 and 124). The method of claim 1 also includes converting the speech to text (page 50, lines 27-31, and Figure 19, elements 554, 552, and 556). The method of claim 1 also includes detecting a total sound level for the user (page 51, lines 1-3, and Figure 19, elements 558, 562, 560, and 560). The method of claim 1 also includes determining whether to display the text in dependence upon the total sound level for the user (page 51, lines 1-7, and Figure 19, elements 558, 562, 560, and 560).

Claim 13 recites a computer-implemented system for differential dynamic content delivery (page 49, lines 27-29, and Figure 19), the system comprising a computer processor, a computer memory operatively coupled to the computer processor, the computer memory having disposed within it computer program instructions. The system of claim 13 includes computer program instructions capable of providing a session document for a presentation (page 49, lines 27-29, and Figure 19, elements 450 and 266), wherein the session document includes a session grammar and a session structured document (page 49, line 29 - page 50, line 3, and Figure 19, elements 266, 258, 256, and 450). The system of claim 13 also includes computer program instructions capable of selecting from the session structured document a classified structural element in dependence upon user classifications of a user participant in the presentation (page 50, lines 5-7, and Figure 19, elements 452, and 465). The system of claim 13 also includes computer program instructions capable of presenting the selected structural element to

the user (page 50, lines 14-15, and Figure 19, elements 454 and 456). The system of claim 13 also includes computer program instructions capable of streaming speech to the user from one or more users participating in the presentation (page 50, lines 24-27, and Figure 19, elements 568 and 124). The system of claim 13 also includes computer program instructions capable of converting the speech to text (page 50, lines 27-31, and Figure 19, elements 554, 552, and 556). The system of claim 13 also includes computer program instructions capable of detecting a total sound level for the user (page 51, lines 1-3, and Figure 19, elements 558, 562, 560, and 560). The system of claim 13 also includes computer program instructions capable of determining whether to display the text in dependence upon the total sound level for the user (page 51, lines 1-7, and Figure 19, elements 558, 562, 560, and 560).

Claim 25 recites a computer program product for differential dynamic content delivery (page 49, lines 27-29, and Figure 19), the computer program product disposed in a signal bearing medium, the computer program product comprising computer program instructions. The computer program product of claim 25 includes computer program instructions capable of providing a session document for a presentation (page 49, lines 27-29, and Figure 19, elements 450 and 266), wherein the session document includes a session grammar and a session structured document (page 49, line 29 - page 50, line 3, and Figure 19, elements 266, 258, 256, and 450). The computer program product of claim 25 also includes computer program instructions capable of selecting from the session structured document a classified structural element in dependence upon user classifications of a user participant in the presentation (page 50, lines 5-7, and Figure 19, elements 452, and 465). The computer program product of claim 25 also includes computer program instructions capable of presenting the selected structural element to the user (page 50, lines 14-15, and Figure 19, elements 454 and 456). The computer program product of claim 25 also includes computer program instructions capable of streaming speech to the user from one or more users participating in the presentation (page 50, lines 24-27, and Figure 19, elements 568 and 124). The computer program product of claim 25 also includes computer program instructions capable of converting the speech to text (page 50, lines 27-31, and Figure 19, elements 554, 552, and 556). The

computer program product of claim 25 also includes computer program instructions capable of detecting a total sound level for the user (page 51, lines 1-3, and Figure 19, elements 558, 562, 560, and 560). The computer program product of claim 25 also includes computer program instructions capable of determining whether to display the text in dependence upon the total sound level for the user (page 51, lines 1-7, and Figure 19, elements 558, 562, 560, and 560).

### GROUND OF REJECTION

In accordance with 37 CFR § 41.37(c)(1)(vi), Appellants provide the following concise statement of each ground of rejection:

1. Claims 25-38 are rejected under 35 U.S.C. § 101 because the claimed invention is directed to non-statutory subject matter.
2. Claims 1-38 are rejected under 35 U.S.C. § 103(a) as being unpatentable over Freire, *et al.* (hereinafter “Freire”), “WebViews: Accessing Personalized Web Content and Services,” *Proceedings of the 10<sup>th</sup> International Conference on World Wide Web*, May 2001, ACM Press, P. 576-586, in view of Brittan, *et al.* (hereinafter “Brittan”), U.S. Patent No. 6,725,199, issued April 2004.

### ARGUMENT

Appellants present the following arguments pursuant to 37 CFR § 41.37(c)(1)(vii) regarding the grounds of rejection in the present case.

**Argument Regarding The First Ground Of Rejection:  
Claims 25-38 Are Rejected Under 35 U.S.C. § 101 Because The  
Claimed Invention Is Directed To Non-Statutory Subject Matter**

Claims 25-38 stand rejected under 35 U.S.C. § 101 because the claimed invention is directed to nonstatutory subject matter. The Office Action states:

In regard to independent claim 25, claim 25 is nonstatutory because the claimed invention claims a computer program product, i.e., the description or expression of a program, and is thus directed toward neither computer components nor statutory processes and do not define any structural or functional interrelationships between the computer program and other claimed elements of a computer which permit the computer program's functionality to be realized. (*Interim Guidelines for Examination of Patent Applications for Patent Subject Matter Eligibility*, hereinafter "Interim Guidelines", p. 53-54).

That is, the Office Action takes the position that the computer program product of claims 25-38 is a description or expression of a program and as such the computer program product of claims 25-38 does not define any structural and functional interrelationship that permits the computer program product's functionality to be realized. Appellants respectfully note in response, however, that MPEP § 2106 states, "...a claimed computer-readable medium encoded with a data structure defines structural and functional interrelationships between the data structure and the computer software and hardware components which permit the data structure's functionality to be realized, and is thus statutory." The computer program product of claims 25-38 is computer readable. The computer program product of claims 25-38 is disposed in a signal bearing medium. The signal bearing medium as claimed in claims 25-38 is therefore computer readable. Examples of a signal bearing medium include a recordable medium (claim 37), and a transmission medium (claim 38). Because the signal bearing medium is computer readable, the computer program product of claims 25-38, disposed within such medium, defines structural and functional interrelationships that permit the computer program product's functionality to be realized. Appellants respectfully submit that claims 25-38 are therefore patentable under 35 U.S.C. § 101, and request reconsideration of claims 25-38.

The Office Action also takes the position that a signal does not fall within one of the four statutory classes of 35 U.S.C. § 101 and as such claims 25-38 are directed to non-statutory subject matter. Appellants do not claim a signal in claims 25-38 but instead claim a computer program product disposed in a signal bearing medium. That is, claims 25 - 38 claim a product - a computer program product. Furthermore, the computer

program product of claims 25-38 is disposed in a medium. Although the medium is described as a 'signal bearing medium,' claims 25-38 do not claim a signal, but instead claim a medium. Claims 25-38 are therefore not directed to a signal, but instead are directed to a computer program product disposed in a signal bearing medium. Appellants respectfully submit that claims 25-38 are therefore patentable under 35 U.S.C. § 101, and request reconsideration of claims 25-38.

**Argument Regarding The Second Ground Of Rejection: Claims 1-38  
Are Rejected Under 35 U.S.C. § 103(A) As Being Unpatentable Over Freire,  
et al. (Hereinafter "Freire"), "Webviews: Accessing Personalized Web Content  
And Services," Proceedings Of The 10th International Conference On World  
Wide Web, May 2001, ACM Press, P. 576-586, In View Of Brittan, et al.  
(Hereinafter "Brittan"), U.S. Patent No. 6,725,199, Issued April 2004**

Claims 1-36 stand rejected under 35 U.S.C § 103(a) as unpatentable over Brittan in view of Freire. Appellants respectfully traverse each rejection. To establish a prima facie case of obviousness, three basic criteria must be met. *Manual of Patent Examining Procedure* §2142. The first element of a prima facie case of obviousness under 35 U.S.C. § 103 is that there must be a suggestion or motivation to combine the references. *In re Vaeck*, 947 F.2d 488, 493, 20 USPQ2d 1438, 1442 (Fed. Cir. 1991). The second element of a prima facie case of obviousness under 35 U.S.C. § 103 is that there must be a reasonable expectation of success in the proposed combination of the references. *In re Merck & Co., Inc.*, 800 F.2d 1091, 1097, 231 USPQ 375, 379 (Fed. Cir. 1986). The third element of a prima facie case of obviousness under 35 U.S.C. § 103 is that the proposed combination of the references must teach or suggest all of Applicants' claim limitations. *In re Royka*, 490 F.2d 981, 985, 180 USPQ 580, 583 (CCPA 1974). As will be shown below, The proposed combination of Brittan and Freire cannot establish a prima facie case of obviousness because the proposed combination does not teach each and every element of the claims of the present application, there is no suggestion or motivation to make the proposed combination, and there is no reasonable expectation of success in the proposed combination.

**The Combination Of Brittan and Freire  
Does Not Teach all Of Appellants' Claim Limitations**

To establish a prima facie case of obviousness, the proposed combination of Brittan and Freire must teach or suggest all of Appellants' claim limitations. *In re Royka*, 490 F.2d 981, 985, 180 USPQ 580, 583 (CCPA 1974). Independent claim 1 of the present application claims:

1. A computer-implemented method for differential dynamic content delivery, the method comprising:  
  
providing a session document for a presentation, wherein the session document includes a session grammar and a session structured document;  
  
selecting from the session structured document a classified structural element in dependence upon user classifications of a user participant in the presentation;  
  
presenting the selected structural element to the user;  
  
streaming speech to the user from one or more users participating in the presentation;  
  
converting the speech to text;  
  
detecting a total sound level for the user; and  
  
determining whether to display the text in dependence upon the total sound level for the user.

**Freire Neither Discloses Nor Suggests Providing A Session  
Document For A Presentation, Wherein The Session Document Includes  
A Session Grammar And A Session Structured Document**

The Office Action takes the position that Freire at page 578, section 2.1, and page 579, figure 2, discloses the first element of claim 1: providing a session document for a presentation, wherein the session document includes a session grammar and a session structured document. Appellants respectfully note in response, however, that what Freire at page 578, section 2.1, in fact discloses is:



To create a Web view, a user must first specify the page to be clipped. If the page requires multiple steps to be retrieved and does not have a well-defined URL, the user can use the recording component of the applet to create the script to access the page. It has a VCR-style interface to transparently record browsing action-users can simply navigate their way to the final page while their actions (links traversed, forms filled along with the user inputs, and any other interactions with active content) are transparently recorded and saved in a *smart bookmark* (SMB).

During recording, if the user is required to fill out forms, she can optionally specify which field values are to be stored in the Web view specification itself, and which ones are to be requested from the user every time the Web view is executed. This allows the user to create parameterized Web views. For example, a Web view to retrieve a restaurant list from the Yellow Pages at mapsonus.com can have a zip code parameter, so the user does not need to create a separate Web view for each city. Also, for security reasons, a user may choose not to save certain kinds of information such passwords inside a Web view (parameterization issues are discussed in Section 2.5), or to save it encrypted. An excerpt of the SMB to retrieve the itinerary page from Travelocity.com (discussed in Example 1.1) is shown in Figure 2.

In addition, what Freire at page 579, figure 2, in fact discloses is:

```
<BOOKMARK id="Juliana_travel">
<URL> http://travelocity.com</URL>
<LINK>
<loc> document.links [8] </loc>
<href> <![CDATA[http://dpsl.travelocity.com/
lonlogin.ctlptr_module=AIRGESEQ=1]]> </href>
<text> null </test> <target> null
</target>
</LINK>
<FORM>
<!--9 Best Itineraries link -->...
</link>
<FORM>
<loc> document.forms[0] </loc>
</action>
<![CDATA [https://dpal.travelocity.com:443/
lognmain.ctl?SEQ=1]]>
</action>
<method> POST </method> <name> null
</name>
<target> null </target>
<ATTRS>
```

```

<ATTR> <name> trip.option </name>
      <loc> 5 </loc>
      <type> radio </type>
      <prop> stored </prop>
      <val> roundtrp </val> </ATTR>
<ATTR> <name> depart.airport </name>
      <loc> 10 </loc>
      <type> text </type>
      <prop> stored </prop>
      <val> EWR </val> </ATTR>
<ATTR> <name> depart.month </name>
      <loc> 11 </loc>
      <type> select-one </type>
      <prop> stored </prop>
      <selected_index> 3 </selected_index>
      <text> Apr </text> </ATTR>
<ATTR> <name> depart_day </name>
      <loc> 12 </loc>
      <type> select-one </type>
      <prop> stored </prop>
      <selected_index> 28 </selected_index>
      <text> 29 </text> </ATTR>
...
</ATTRS>
</FORM>
</BOOKMARK>

```

That is, Freire at page 578, section 2.1 and page 579, figure 2, discloses creating a web view where a user's actions are transparently recorded and saved in a smart bookmark. Freire's creating a web view where a user's actions are transparently recorded and saved in a smart bookmark does not disclose providing a session document for a presentation, wherein the session document includes a session grammar and a session structured document as claimed in the present application. Freire's 'smart bookmark' is a markup language document representing a 'web view.' Freire's 'web view' does not disclose or suggest a session document as claimed in the present application because a session document as claimed in the present application is for a presentation. The 'presentation' as claimed in the present application is a multimedia presentation with a presenter and a number of user participants. Freire's 'web view' is not concerned at all with a multimedia presentation with a presenter and a number of user participants. Freire's 'web view' is a recording of a user's actions while the user navigates within a web

browser to a final webpage. Furthermore, Freire only discloses recording a single user's actions to create a single user's web view, with no mention whatsoever of a presenter and a number of user participants. Because Freire's 'web view' does not disclose or suggest a session document for a presentation as claimed in the present application it cannot be said that Freire discloses or suggests providing a session document for a presentation, wherein the session document includes a session grammar and a session structured document as claimed in the present application. The Office Action therefore cannot establish a prima facie case of obviousness. The rejections of claims 1-38 should be withdrawn, and the claims should be allowed.

**Freire Neither Discloses Nor Suggests Selecting  
From The Session Structured Document A Classified  
Structural Element In Dependence Upon User Classifications  
Of A User Participant In The Presentation; Presenting  
The Selected Structural Element To The User**

The Office Action takes the position that Freire at pages 582-583, section 3.2, discloses the second and third elements of claim 1: selecting from the session structured document a classified structural element in dependence upon user classifications of a user participant in the presentation and presenting the selected structural element to the user. Appellants respectfully note in response, however, that what Freire at pages 582-583, section 3.2, in fact discloses is:

Section 2 assumed the presence of transcoding proxies to enable access to various devices. Note, however, that the transcoding functionality can also be incorporated into the WebViews server, with the added advantage that the user can now annotate the Web view and supply extra information that can be used in the transcoding process to produce better quality content, and a more user-friendly experience. The main drawback of this tight-coupling is that the transcoding engine and server must agree on a set of annotations. If they are developed by separate parties, this will sometimes discourage such an architecture, unless, of course, a standard for such annotations exists.

We have built an engine that transcodes clipped HTML content into VoiceXML. In what follows, we describe its architecture, and discuss how it can be combined with the WebViews server to create VoiceViews. Note that we still assume the presence of a Voice gateway (running a

VoiceXML interpreter) that connects the PSTN network to the IP network. The user calls a phone number, and the VoiceXML interpreter requests the WebViews server for any VoiceXML dialogs to be played to the user.

The current transcoding architecture is diagramed in Figure 5. The usage scenario is as follows. When the user calls a special phone number, a fixed caller identification VoiceXML dialog is started. The dialogue attempts to identify the user using his caller ID (which it uses as userid); if that is not available, the system interrogates the user for the userid. Once the userid is obtained, the list of Web views associated with that user is looked up (via GenerateVoiceViewList) from the User Profile DB, and a VoiceXML dialog is generated that prompts the user to select between one of his recorded Web views. The user can then make his selection via touch-tone or spoken input. Once the user makes a choice, the VoiceXML interpreter passes the Web view information to the Transcoding Engine, which in turn queries the WebViews execution engine to execute the Web view. The Transcoding Engine converts the clipped content into VoiceXML, utilizing the extra annotations supplied with the Web view.

Figure 6 shows the VoiceXML dialogue generated by our transcoder for the clipping (appropriately tidied) of the Yahoo! car page shown in Figure 4. In the dialogue, each car listing is transcoded as a field of a VoiceXML form. The form contains all the data transcoded from a single top-level table. The transcoded output could also have been output as a single block so that the entire contents would have been read as a unit using TTS, but that would have given the user no option but to listen to the whole table being read or to hang up. As it is, the script listens for special keywords, namely ``next" and ``skip", allowing the user to quit hearing details of a single row or of the rest of the table. If the user says nothing (noinput) or something incomprehensible (nomatch), the script goes to the next row.

That is, Freire at pages 582-583, section 3.2, discloses that transcoding functionality can be on a WebViews server as well as a transcoding proxy, that the authors of Freire have built an HTML to VoiceXML transcoder, the transcoder's architecture, a usage example for the transcoder, and a VoiceXML dialogue generated by the transcoder from an HTML document. Freire's transcoding functionality can be on a WebViews server as well as a transcoding proxy, that the authors of Freire have built an HTML to VoiceXML transcoder, the transcoder's architecture, a usage example for the transcoder, and a VoiceXML dialogue generated by the transcoder from an HTML document does not disclose selecting from the session structured document a classified structural element in dependence upon user classifications of a user participant in the presentation as claimed

in the present application. At no point does Freire disclose or even suggest a ‘user classification’ as claimed in the present application. The ‘user classification’ may be a characteristic of a user such as the user’s department, the user’s area and level of technical expertise, the user’s security clearance level, and so on. The selected structural element presented to a user in the third element of claim 1 in the present application is a selected structural element that is selected in dependence upon such a user classification of the user participant to whom the content is presented, including, for example, the user’s department, security clearance, expertise, and so on. Freire does not disclose, suggest, or even mention selecting anything from a structured document in dependence upon such a user classification. Freire’s transcoding functionality on a WebViews server, HTML to VoiceXML transcoder, the transcoder’s architecture, transcoder usage example, and example VoiceXML dialogue generated by the transcoder from an HTML document neither discloses nor suggests selecting from a session structured document a classified structural element in dependence upon user classifications of a user participant in the presentation claimed in the present application.

Freire does not, and in fact, cannot disclose or suggest presenting the selected structural element to the user as claimed in the present application. The ‘structural element’ that is selected as claimed in the present application, is selected in dependence upon ‘user classifications of a user participant.’ As explained in detail above, Freire does not disclose or suggest ‘user classifications of a user participant’ as claimed in the present application. Freire cannot disclose therefore presenting the selected structural element to the user as claimed in the present application. Because the proposed combination of references does not disclose or suggest each and every element and limitation of Appellants’ claims, the proposed combination does not render Appellants’ claims obvious. The Office Action therefore cannot establish a prima facie case of obviousness, and the obviousness rejections of claims 1-36 should be withdrawn and the claims allowed.

**Brittan Neither Discloses Nor Suggests Converting  
The Speech To Text; Detecting A Total Sound Level For  
The User; And Determining Whether To Display The Text In  
Dependence Upon The Total Sound Level For The User**

The Office Action takes the position that Brittan at column 8, line 34 - column 9, line 26, and column 9 lines, 32-38, discloses the fifth, sixth, and seventh elements of claim 1: converting the speech to text, detecting a total sound level for the user, and determining whether to display the text in dependence upon the total sound level for the user.

Appellants respectfully note in response, however, that what Brittan at column 8, line 34 - column 9, line 26, and column 9 lines, 32-38, in fact discloses is:

Changing dialog style can also be effected for other reasons concerning the intelligibility of the speech heard by the user. Thus, if the user is in a noisy environment (for example, in a vehicle) then the system can be arranged to narrow and direct the dialogue, reducing the chance of misunderstanding. On the other hand, if the environment is quiet, the dialogue could be opened up, allowing for mixed initiative. To this end, the speech system is provided with a background analysis block 45 connected to sound input source 16 in order to analyze the input sound to determine whether the background is a noisy one; the output from block 45 is fed to the style selection block 46 to indicate to the latter whether background is noisy or quiet. It will be appreciated that the output of block 45 can be more fine grain than just two states. The task of the background analysis block 45 can be facilitated by (i) having the TTS 6 inform it when the latter is outputting speech (this avoids feedback of the sound output being misinterpreted as noise), and (ii) having the speech recognizer 5 inform the block 45 when the input is recognizable user input and therefore not background noise (appropriate account being taken of the delay inherent in the recognizer determining input to be speech input).

Where both intelligibility as measured by the confidence score output by the classifier and the level background noise are used to effect the selected dialog style, it may be preferable to feed the confidence score directly to the style selection block 45 to enable block 45 to use this score in combination with the background-noise measure to determine which style to set.

It is also possible to provide for user selection of dialog style.

Multi-modal output (FIG. 6)--more and more devices, such as third generation mobile appliances, are being provided with the means for

conveying a concept using both voice and a graphical display. If confidence is low in the synthesized speech, then more emphasis can be placed on the visual display of the concept. For example, where a user is receiving travel directions with specific instructions being given by speech and a map being displayed, then if the classifier produces a low confidence score in relation to an utterance including a particular street name, that name can be displayed in large text on the display. In another scenario, the display is only used when clarification of the speech channel is required. In both cases, the display acts as a supplementary modality for clarifying or exemplifying the speech channel. FIG. 6 illustrates an implementation of such an arrangement in the case of a generalized supplementary modality (whilst a visual output is likely to be the best form of supplementary modality in most cases, other modalities are possible such as touch/feel-dependent modalities). In FIG. 6, the language generator 23 provides not only a text output to the TTS 6 but also a supplementary modality output that is held in buffer 48. This supplementary modality output is only used if the output of the classifier 41 indicates a low confidence in the current speech output; in this event, the CAC causes the supplementary modality output to be fed to the output constructor 28 where it is converted into a suitable form (for example, for display). In this embodiment, the speech output is always produced and, accordingly, the speech output buffer 44 is not required.

...

The supplementary modality can, in fact, be used as an alternative modality--that is, it substitutes for the speech output for a particular utterance rather than supplementing it. In this case, the speech output buffer 44 is retained and the CAC 43 not only controls output from the supplementary-modality output buffer 48 but also controls output from buffer 44 (in anti-phase to output from buffer 48).

That is, Brittan at this reference point discloses methods of changing dialog styles for speech output from text-to-speech conversions. Brittan's methods of changing dialog styles for speech output from text-to-speech conversions do not disclose or suggest converting the speech to text as claimed in the present application. 'The speech' that is converted to text as claimed in the present application is speech that is streamed to the user from one or more users participating in the presentation. All the speech delivered in Brittan is delivered directly from a speech synthesis engine to a user, not from one user to another user. Brittan has no concern, no teaching or suggestion, regarding multimedia presentations to multiple users nor with streaming speech to one user from other users

participating in a presentation as claimed in the fourth element of the present application: “streaming speech to the user from one or more users participating in the presentation.” Because Brittan does not disclose or suggest “streaming speech” as claimed in the present application it cannot be said the Brittan discloses or suggests converting such speech to text as claimed in the present application.

Brittan’s methods of changing dialog styles for speech output from text-to-speech conversions do not disclose or suggest detecting a total sound level for the user. According to Brittan, the changes in dialog style may be made according to background noise levels or according to user selection. Brittan also discloses that supplemental modalities such as graphical displays may be used in addition to speech output or as a substitute for speech output. Brittan does not disclose or even suggest however that such supplemental modalities may be changed according to the Brittan’s background noise levels. ‘The total sound level for the user’ as claimed in the present application however, is in fact used to determine whether to display the text. That is, Brittan does not disclose or suggest making any determination whether to display text (Brittan’s supplemental modalities) in dependence upon the total sound level for the user as claimed in the present application. Brittan does not disclose or suggest therefore detecting the total sound level for the user as claimed in the present application, where the total sound level is used to determine whether to display the text.

In addition to the fact that Brittan does not disclose or suggest making any determination whether to display text (Brittan’s supplemental modalities) in dependence upon the total sound level for the user as claimed in the present application, there is a second reason that Brittan does not disclose or suggest detecting the total sound level as claimed in the present application: Brittan cannot disclose or suggest determining whether to display *the text* in dependence upon the total sound level for the user. As explained above, Brittan does not disclose or suggest converting the speech to text and as such cannot be said to disclose or suggest determining whether to display such text. Because Brittan does not disclose or suggest determining whether to display the text and because such determination is made in dependence upon the total sound level for the user, it cannot be



said that Brittan discloses or suggests detecting the total sound level for the user as claimed in the present application. Brittan therefore does not disclose or suggest the fifth, sixth, or seventh elements of claim 1 of the present application. The Office Action therefore cannot establish a prima facie case of obviousness, the rejections of claims 1-36 should be withdrawn, and the claims allowed.

**No Suggestion or Motivation to Modify Brittan**

To establish a prima facie case of obviousness, there must be a suggestion or motivation to modify Brittan. *In re Vaeck*, 947 F.2d 488, 493, 20 USPQ2d 1438, 1442 (Fed. Cir. 1991). The suggestion or motivation to modify Brittan must come from the teaching of the references themselves, and the Examiner must explicitly point to the teaching within the references, or to knowledge of those of skill in the art, suggesting the proposed modification. Absent such a showing, the Examiner has impermissibly used “hindsight” occasioned by Appellants’ own teaching to reject the claims. *In re Surko*, 11 F.3d 887, 42 U.S.P.Q.2d 1476 (Fed. Cir. 1997); *In re Vaeck*, 947 F.2d 488m 20 U.S.P.Q.2d 1438 (Fed. Cir. 1991); *In re Gorman*, 933 F.2d 982, 986, 18 U.S.P.Q.2d 1885, 1888 (Fed. Cir. 1991); *In re Bond*, 910 F.2d 831, 15 U.S.P.Q.2d 1566 (Fed. Cir. 1990); *In re Laskowski*, 871 F.,2d 115, 117, 10 U.S.P.Q.2d 1397, 1398 (Fed. Cir. 1989).

There can be no suggestion or motivation to combine in this case because neither reference is concerned with differential content delivery in dependence upon user characteristics. Moreover, it would never occur to a person of skill in the art to combine the references to achieve differential content delivery in dependence upon user characteristics because combining them would not achieve differential content delivery in dependence upon user characteristics. Neither reference is concerned in any way with differential delivery of content based upon user characteristics. Neither Brittan nor Freire discloses or suggests therefore such claim elements as presenting the selected structural element to the user, streaming speech to the user from one or more users participating in the presentation, or determining whether to display the text in dependence upon the total sound level for the user as claimed in the present application. Brittan only discloses a

speech synthesizer with two or more synthesis engines providing various styles of dialog, with a dialog style selected according to quality of speech form utterance and the like. Brittan makes no mention or suggestion whatsoever of tailoring content according to characteristics or classifications of users. Freire is only concerned with easing web site access. Freire is not concerned with multimedia presentations and Freire neither teaches nor suggests tailoring content according to characteristics or classifications of users. What the combination of Brittan and Freire then actually discloses is simplified web site access using a speech synthesizer with two or more synthesis engines providing various styles of dialog. The combination does not disclose in any way anything having to do with multimedia presentations or classifications of users as claimed in the present application. That is, neither reference discloses or suggests any aspect of differential content delivery in dependence upon user characteristics as claimed in the present application. It would never occur then to a person of skill in the art to combine the two references to attempt to produce such differential content delivery in dependence upon user characteristics.

**Brittan And Freire Are Non-Analogous Art And Therefore Are  
Unavailable As References Against The Present Application**

The field of endeavor in the present application is generally differential dynamic content delivery, and particularly differential dynamic content delivery for multimedia presentations based upon classifications of users with text display in dependence upon sound level. Brittan is about a speech synthesizer with two or more synthesis engines providing various styles of dialog, with a dialog style selected according to quality of speech form utterance and the like. Freire is concerned with easing web site access. An inventor concerned with differential dynamic content delivery for multimedia presentations based upon classifications of users would not be reasonably expected to examine the art for references dealing with speech synthesizer dialog selection or easing web site access. Both Brittan and Freire therefore are non-analogous and not available as references against the present application under 35 U.S.C. § 103.

**Relations Among Claims**

Independent claim 1 claims method for differential dynamic content delivery according to embodiments of the present invention. Independent claims 13 and 25 respectively claim system and computer program product aspects of differential dynamic content delivery according to embodiments of the present invention. Claim 1 is allowable for the reasons set forth above. Claims 13 and 25 are allowable because claim 1 is allowable. The rejections of claims 13 and 25 therefore should be withdrawn, and claims 13 and 25 should be allowed.

Claims 2-12, 14-24, and 26-36 depend respectively from independent claims 1, 13, and 25. Each dependent claim includes all of the limitations of the independent claim from which it depends. Because the combination of Brittan and Freire does not establish a prima facie case of obviousness for the independent claims, so also the combination of Brittan and Freire does not establish a prima facie case of obviousness for any dependent claim. The rejections of Claims 2-12, 14-24, and 26-36 therefore should be withdrawn, and these claims also should be allowed.

**Conclusion of Appellants' Arguments**

Claims 1-36 stand rejected for obviousness under 35 U.S.C. § 103(a) as being unpatentable over Freire in view of Brittan. For the reasons set forth above, however, the proposed combination of Freire and Brittan fails to establish a prima facie case of obviousness. The rejection of claims 1-36 should therefore be withdrawn, and the claims should be allowed. Reconsideration of claims 1-36 in light of the present remarks is respectfully requested.

**APPEAL BRIEF**

The Commissioner is hereby authorized to charge or credit Deposit Account No. 09-0447 for any fees required or overpaid.

Respectfully submitted,

Date: January 26, 2007

By: 

Thomas Fortenberry  
Reg. No. 56,537  
Biggers & Ohanian, LLP  
P.O. Box 1469  
Austin, Texas 78767-1469  
Tel. (512) 472-9881  
Fax (512) 472-9887  
ATTORNEY FOR APPELLANTS

**APPENDIX OF CLAIMS**  
**ON APPEAL IN PATENT APPLICATION OF**  
**WILLIAM K BODIN, *ET AL.*, SERIAL NO. 10/756,159**

CLAIMS

What is claimed is:

1. A computer-implemented method for differential dynamic content delivery, the method comprising:  
  
providing a session document for a presentation, wherein the session document includes a session grammar and a session structured document;  
  
selecting from the session structured document a classified structural element in dependence upon user classifications of a user participant in the presentation;  
  
presenting the selected structural element to the user;  
  
streaming speech to the user from one or more users participating in the presentation;  
  
converting the speech to text;  
  
detecting a total sound level for the user; and  
  
determining whether to display the text in dependence upon the total sound level for the user.
2. The method of claim 1 wherein the total sound level for the user includes ambient noise and the method includes detecting an ambient noise level for the user.

3. The method of claim 2 wherein detecting an ambient noise level for the user further comprises temporarily interrupting the speech streaming to the user and measuring a sound level on the user's voice channel during the interruption and while the user is not speaking.
4. The method of claim 2 wherein determining whether to display the text in dependence upon the total sound level for the user further comprises determining to display the text to the user if the ambient noise level is above a predetermined threshold.
5. The method of claim 1 wherein:  
  
the total sound level comprises the streaming speech plus ambient noise; and  
  
determining whether to display the text further comprises determining whether to display the text in dependence upon a ratio of the total sound level to the ambient noise level.
6. The method of claim 5 further comprising displaying the text if the ratio of the total sound level to the ambient noise level is less than a predetermined minimum.
7. The method of claim 1 wherein selecting a classified structural element further comprises selecting a classified structural element having an associated classification identifier that corresponds to the user classification.
8. The method of claim 1 further comprising creating a session document from a presentation document, including:

identifying a presentation document for a presentation, the presentation document including a presentation grammar and a structured document having structural elements classified with classification identifiers;

identifying a user participant for the presentation, the user having a user profile comprising user classifications; and

filtering the structured document in dependence upon the user classifications and the classification identifiers.

9. The method of claim 8 further comprising filtering the presentation grammar, in dependence upon the extracted structural elements, into a session grammar in the session document.

10. The method of claim 8 further comprising creating a presentation document, including:

creating, in dependence upon an original document, a structured document comprising one or more structural elements;

classifying a structural element of the structured document according to a presentation attribute; and

creating a presentation grammar for the structured document, wherein the presentation grammar for the structured document includes grammar elements each of which includes an identifier for at least one structural element of the structured document.

11. The method of claim 10 wherein classifying a structural element comprises:

identifying a presentation attribute for the structural element;

identifying a classification identifier in dependence upon the presentation attribute; and

inserting the classification identifier in association with the structural element in the structured document.

12. The method of claim 10 wherein creating a presentation grammar for the structured document comprises:

identifying the content type of the original document;

selecting, in dependence upon the content type, a full presentation grammar from among a multiplicity of full presentation grammars; and

filtering the full presentation grammar into a presentation grammar for the structured document in dependence upon the structural elements of the structured document.

13. A system for differential dynamic content delivery, the system comprising a computer processor, a computer memory operatively coupled to the computer processor, the computer memory having disposed within it computer program instructions capable of:

providing a session document for a presentation, wherein the session document includes a session grammar and a session structured document;

selecting from the session structured document a classified structural element in dependence upon user classifications of a user participant in the presentation;

presenting the selected structural element to the user;



streaming speech to the user from one or more users participating in the presentation;

converting the speech to text;

detecting a total sound level for the user; and

determining whether to display the text in dependence upon the total sound level for the user.

14. The system of claim 13 wherein the total sound level for the user includes ambient noise and the system includes computer program instructions capable of detecting an ambient noise level for the user.
15. The system of claim 14 wherein detecting an ambient noise level for the user further comprises temporarily interrupting the speech streaming to the user and measuring a sound level on the user's voice channel during the interruption and while the user is not speaking.
16. The system of claim 14 wherein determining whether to display the text in dependence upon the total sound level for the user further comprises determining to display the text to the user if the ambient noise level is above a predetermined threshold.
17. The system of claim 13 wherein:  
  
the total sound level comprises the streaming speech plus ambient noise; and

determining whether to display the text further comprises determining whether to display the text in dependence upon a ratio of the total sound level to the ambient noise level.

18. The system of claim 17 further comprising computer program instructions capable of displaying the text if the ratio of the total sound level to the ambient noise level is less than a predetermined minimum.

19. The system of claim 13 wherein selecting a classified structural element further comprises selecting a classified structural element having an associated classification identifier that corresponds to the user classification.

20. The system of claim 13 further comprising computer program instructions capable of creating a session document from a presentation document, including:

identifying a presentation document for a presentation, the presentation document including a presentation grammar and a structured document having structural elements classified with classification identifiers;

identifying a user participant for the presentation, the user having a user profile comprising user classifications; and

filtering the structured document in dependence upon the user classifications and the classification identifiers.

21. The system of claim 20 further comprising computer program instructions capable of filtering the presentation grammar, in dependence upon the extracted structural elements, into a session grammar in the session document.

22. The system of claim 20 further comprising computer program instructions capable of creating a presentation document, including:

creating, in dependence upon an original document, a structured document comprising one or more structural elements;

classifying a structural element of the structured document according to a presentation attribute; and

creating a presentation grammar for the structured document, wherein the presentation grammar for the structured document includes grammar elements each of which includes an identifier for at least one structural element of the structured document.

23. The system of claim 22 wherein classifying a structural element comprises:

identifying a presentation attribute for the structural element;

identifying a classification identifier in dependence upon the presentation attribute; and

inserting the classification identifier in association with the structural element in the structured document.

24. The system of claim 22 wherein means for creating a presentation grammar for the structured document comprises:

identifying the content type of the original document;

selecting, in dependence upon the content type, a full presentation grammar from among a multiplicity of full presentation grammars; and

filtering the full presentation grammar into a presentation grammar for the structured document in dependence upon the structural elements of the structured document.

25. A computer program product for differential dynamic content delivery, the computer program product the computer program product disposed in a signal bearing medium, the computer program product comprising computer program instructions capable of:

providing a session document for a presentation, wherein the session document includes a session grammar and a session structured document;

selecting from the session structured document a classified structural element in dependence upon user classifications of a user participant in the presentation;

presenting the selected structural element to the user;

streaming speech to the user from one or more users participating in the presentation;

converting the speech to text;

detecting a total sound level for the user; and

determining whether to display the text in dependence upon the total sound level for the user.

26. The computer program product of claim 25 wherein the total sound level for the user includes ambient noise and the computer program product includes computer program instructions capable of detecting an ambient noise level for the user.

27. The computer program product of claim 26 wherein detecting an ambient noise level for the user further comprises temporarily interrupting the speech streaming to the user and measuring a sound level on the user's voice channel during the interruption and while the user is not speaking.
28. The computer program product of claim 26 wherein determining whether to display the text in dependence upon the total sound level for the user further comprises determining to display the text to the user if the ambient noise level is above a predetermined threshold.
29. The computer program product of claim 25 wherein:  
  
the total sound level comprises the streaming speech plus ambient noise; and  
  
determining whether to display the text further comprises determining whether to display the text in dependence upon a ratio of the total sound level to the ambient noise level.
30. The computer program product of claim 27 further comprising computer program instructions capable of displaying the text if the ratio of the total sound level to the ambient noise level is less than a predetermined minimum.
31. The computer program product of claim 25 wherein selecting a classified structural element further comprises selecting a classified structural element having an associated classification identifier that corresponds to the user classification.
32. The computer program product of claim 25 further comprising computer program instructions capable of creating a session document from a presentation document, including:

identifying a presentation document for a presentation, the presentation document including a presentation grammar and a structured document having structural elements classified with classification identifiers;

identifying a user participant for the presentation, the user having a user profile comprising user classifications; and

filtering the structured document in dependence upon the user classifications and the classification identifiers.

33. The computer program product of claim 32 further comprising computer program instructions capable of filtering the presentation grammar, in dependence upon the extracted structural elements, into a session grammar in the session document.

34. The computer program product of claim 32 further comprising computer program instructions capable of creating a presentation document, including:

creating, in dependence upon an original document, a structured document comprising one or more structural elements;

classifying a structural element of the structured document according to a presentation attribute; and

creating a presentation grammar for the structured document, wherein the presentation grammar for the structured document includes grammar elements each of which includes an identifier for at least one structural element of the structured document.

35. The computer program product of claim 34 wherein classifying a structural element comprises:

identifying a presentation attribute for the structural element;

identifying a classification identifier in dependence upon the presentation attribute; and

inserting the classification identifier in association with the structural element in the structured document.

36. The computer program product of claim 34 wherein creating a presentation grammar for the structured document comprises:

identifying the content type of the original document;

selecting, in dependence upon the content type, a full presentation grammar from among a multiplicity of full presentation grammars; and

filtering the full presentation grammar into a presentation grammar for the structured document in dependence upon the structural elements of the structured document.

37. The computer program product of claim 25 wherein the signal bearing medium comprises a recordable medium.

38. The computer program product of claim 25 wherein the signal bearing medium comprises a transmission medium.

**APPENDIX OF EVIDENCE**  
**ON APPEAL IN PATENT APPLICATION OF**  
**WILLIAM K BODIN, *ET AL.*, SERIAL NO. 10/756,159**

This is an evidence appendix in accordance with 37 CFR § 41.37(c)(1)(ix).

There is in this case no evidence submitted pursuant to 37 CFR §§ 1.130, 1.131, or 1.132, nor is there in this case any other evidence entered by the examiner and relied upon by the appellants.



**RELATED PROCEEDINGS APPENDIX**

This is a related proceedings appendix in accordance with 37 CFR § 41.37(c)(1)(x).

There are no decisions rendered by a court or the Board in any proceeding identified pursuant to 37 CFR § 41.37(c)(1)(ii).